549,603

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 November 2004 (18.11.2004)

PCT

(10) International Publication Number
WO 2004/099370 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number:
PCT/US2004/013257

(22) International Filing Date: 28 April 2004 (28.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/467,315 30 April 2003 (30.04.2003) US

(71) Applicant *(for all designated States except US)*: GENECOR INTERNATIONAL, INC. [US/US]; 925 Page Mill Road, Palo Alto, California 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: JONES, Brian, E. [GB/NL]; Gravin Juliana van Stolberglaan 24, NL-2263VA Leidschendam (NL). GRANT, William , D. [GB/GB]; 15 The Tithings, Kibworth Beauchamp Leicestershire LE8 OPU (GB). HEAPHY, Shaun [GB/GB]; 426 Uppingham Road, Leicester Leicestershire LE52DP (GB). GRANT, Susan [GB/GB]; 15 The Tithings, Kibworth Beauchamp Leicestershire LE8 OPU (GB). REES, Helen [GB/GB]; Flat 3, 15 Moor Oaks Road, Sheffield Leicestershire S10 1BX (GB).

(74) Agent: BOYD, Victoria, L.; GENENCOR INTERNATIONAL, INC., 925 Page Mill Road, Palo Alto, California 94304 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,

[Continued on next page]

(54) Title: NOVEL BACILLUS mHKcel CELLULASE

(57) Abstract: The present invention provides a novel cellulase nucleic acid sequence, designated mHKcel, and the corresponding mHKcel amino acid sequence. The invention also provides expression vectors and host cells comprising a nucleic acid sequence encoding mHKcel, recombinant mHKcel proteins and methods for producing the same.

ORF Nucleotide sequence of mHKcel cellulase gene

```
ATGGGTTATA CCCAAGCTAA GTGTATGGTG AAAAAAACGG TCTTGTTTGG      50
TTTAATTCTC TGTTTAGGTG TGTCAATGTT TGTACCAGTT ACATCAGCTG     100
AAGATAGGGT CTCTTCGTCA CAGGTGGATA TCCAATCATA TGTAGCAGAT     150
ATGCAACCTG GCTGGAATTT AGGTAATACA TTTGATGCGA TAGGAGATGA     200
TGAAACAGCA TGGGGAAACC CTCGTGTAAC GAGAGAATTA ATAGAAATGA     250
TTGCTGATGA AGGGTATAAA AGTATTCGTA TCCCAGTCAC ATGGCAAAAT     300
CAAATGGGTG GTTCTCCAGA TTATACAATT AATGAAGATT ATATCAAGCG     350
GGTAGAGCAA GTGATAGATT GGGCGTTGGA GGAAGACTTG TATGTGATGT     400
TAAATGTGCA TCATGACTCA TGGCTGTGGA TGTATGATAT GGAACATAAC     450
TATGATGAGG TGATGGCAAG ATATACAGCT ATTTGGGAAC AATTGTCGGA     500
AAAATTCAAA AACCACTCCC ATAAGTTGAT GTTTGAGAGT GTCAATGAGC     550
CTAGGTTTAC GCAGGAGTGG GGAGAGATTC AAGAAAATCA TCATGCTTAC     600
TTAGAAGATT TAAATAAGAC GTTCTATTAT ATTGTCAGAG AGTCAGGAGG     650
CAATAATGTG GAGCGCCCTT TAGTATTGCC TACGATAGAA ACAGCCACGT     700
CTCAGGATTT ACTAGATCGC TTGTATCAAA CAATGGAAGA CTTGGATGAC     750
CCTCATTTAA TTGCCACGGT TCATTATTAT GGCTTTTGGC CCTTTAGTGT     800
CAATATAGCA GGGTACACCC GTTTTGAACA GGAGACACAA CAAGATATTA     850
TAGACACGTT TGACCGTGTT CATAACACAT TTACAGCGAA TGGGATCCCA     900
GTTGTATTAG GTGAATTTGG TTTGTTAGGC TTTGATAAAA GTACGGACGT     950
CATTCAGCAA GGTGAGAAAT TAAAATTTTT TGAGTTTCTC ATCCATCATC    1000
TCAATGAACG TGATATAACC CATATGTTAT GGGATAACGG TCAGCATTTA    1050
AAGCGAGAAA CTTATTCATG GTATGATCAG GAATTTCATG ACATATTAAA    1100
AGCGAGTTGG GAGGGGCGTT CTGCTACAGC TGAGTCTAAT TTCATTCATG    1150
TGAAGGACGG AGAGCCAATT AGAGATCAAC ATATACAGCT TTACTTAAAC    1200
GGAAATGAGC TAACTGCCCT ACAGGCAGGG GACGAATCGC TTGTACTAGG    1250
AGAGGATTAT GAGCTAGCAG GAGACGTATT AACGCTAAAA GCGGGCATCC    1300
TCACAAGATT AATTACCCCT GGCCAATTAG GAACGAATGC GGTCATCACA    1350
GCTCAATTTA ATTCTGGAGC AGACTGGCGT TTTCAATTAC AGAATGTGGA    1400
CGTGCCAACA GTCGAAAATA CAGATGGCTC AATATGGCAT TTTGCGATCC    1450
CTACCCATTT TAATGGTGAT AGTCTTGCGA CGATGGAAGC TGTTTATGCA    1500
AACGGAGAAT ATGCTGGCCC GCAAGATTGG ACGTCATTTA AAGAATTTGG    1550
CGAGGCGTTT TCCCCTAATT ACGCCACAGG GGAAATTATT ATAACAGAAG    1600
CCTTCTTTAA CGCGGTACGG GATGATGATA TCCATTTAAC ATTTCATTAT    1650
TGGAGCGGAG AGACGGTGGA ATATACATTA CGTAAAAATG GAAATTATGT    1700
TCAAGGTAGA CGGTAA                                         1715
```

WO 2004/099370 A2